

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 29 日現在

機関番号：12601

研究種目：基盤研究(S)

研究期間：2008～2013

課題番号：20222001

研究課題名(和文) 史料デジタル収集の体系化に基づく歴史オントロジー構築の研究

研究課題名(英文) A Study of Constructing of the Historical Ontology with Digitizing of Historical Materials

研究代表者

林 譲 (HAYASHI, Yuzuru)

東京大学・史料編纂所・教授

研究者番号：00164971

交付決定額(研究期間全体)：(直接経費) 151,900,000円、(間接経費) 45,570,000円

研究成果の概要(和文)：東京大学史料編纂所が60年間にわたって収集・蓄積した探訪史料マイクロフィルムをデジタル化し、オープンデジタルによる収集の仕様を確立し、一点目録情報などのメタデータを付与したデジタルデータを格納するアーカイヴハブ(デジタル画像史料収蔵庫)を構築し公開した。あわせて、デジタル画像史料群に基づく先端的プロジェクト・歴史オントロジー構築の研究を推進し、研究成果を公開した。

研究成果の概要(英文)：This Project realized a system named as "ArchiveHub" [archival server for digital images] on the intranet of Historiographical Institute of the University of Tokyo, which has a controlling system for the metadata of these images. As a result the digital images of historical materials in Japan, not only converted from the microfilms collected and accumulated for 60 years by the Institute, but also taken by digital camera, are loaded to this server. Simultaneously this project promoted the advanced study of the historical ontology. Its prototype system is available on the above-mentioned intranet.

研究分野：人文学

科研費の分科・細目：史学・日本史

キーワード：日本史 史料 デジタル アーカイヴハブ 歴史知識学 オントロジー

1. 研究開始当初の背景

(1)研究代表者らが所属する東京大学史料編纂所では、国内外の日本史史料を組織的・系統的に調査・収集し、複成本(影写本・謄写本など)として利用されてきた。なかでも、マイクロフィルム・カメラ撮影による複成は「写真帳」として広く利用されてきたが、2008年以降、複写用印画紙、史料撮影に適した35ミリロールフィルムの生産が中止になり、マイクロフィルムとカメラによる史料採訪・複成作成システムの継続が全く困難な状況に陥った。

(2)一方、史料編纂所では、所蔵する影写本・謄写本・史料稿本類のデジタル化(総計530万件)に取り組み、研究資源の高度情報化を進めていた。

2. 研究の目的

本研究は、これらの成果・資源を継承し、アナログ媒体のデジタル仕様による情報資源化を目指し、史料収集システムのデジタル仕様への再構築を図るものである。その目的とするところは、第一に、採訪史料フィルム類のデジタル化とアーカイヴハブ(デジタル画像史料収蔵庫)の構築、およびデジタル撮影(ボーンデジタル)による史料収集の仕様の確立、第二に、デジタル画像史料群に基づく先端的プロジェクト・歴史オントロジー構築の研究、にある。

3. 研究の方法

(1)史料編纂所が60年間にわたって系統的に蓄積するフィルムベースの史料画像(約490万コマ)をデジタル化する。フィルムの劣化や原本史料の欠如・公開要請の高いものから開始する。

(2)デジタル撮影(ボーンデジタル)による史料収集の仕方を研究し、モデル的な史料収集を推進する。デジタル化した史料画像をファイルサーバに格納し、史料の原秩序と階層構造を反映させたメタデータを付与する。必要に応じて一点ごとの目録情報も

作成し、サーバ空間内にアーカイヴハブ(デジタル画像史料収蔵庫)を構築し、研究資源の高度情報化と共同利用をはかる。

(3)アーカイヴハブのデジタル画像史料群を基礎に、画像研究とテキスト研究における先端的プロジェクトを推進し、史料編纂所の歴史情報処理システム(SHIPS)データベースとアーカイヴハブとの統合を目指す。

(4)歴史知識を蓄積する歴史オントロジーの構築について研究する。

4. 研究成果

(1)本研究を時間軸に沿って整理すると、2008年度は課題の遂行に最適なサーバなどの基盤を整備してデータ蓄積を開始し、2009年度は画像ビューワ・検索システム構築による閲覧・検索可能な環境を創り出し、2010年度はWebコンテンツ保護機能・Hi-Cat Plusなど公開とセキュリティ対策のシステム開発を行ってきた。これらを前提に、蓄積・検索・閲覧・公開の環境が整った2011年度において、本所閲覧室において、マイクロフィルム簡易検索・Hi-CAT Plusなどのシステムによる採訪マイクロフィルムデジタル画像の公開を開始し、最終年度の2012年度において、デジタル化・目録作成・データアップロードなどの諸事業の遂行と改訂作業を進め、秋田県公文書館等との連携システムを実現し、新たな検索方法としての史料所在オントロジーシステムを開発・実装した。年度を繰り越し、研究成果報告書を刊行した。

(2)以下、具体的な成果を報告する。

採訪史料マイクロフィルム(国内)のほぼ全部にあたる418万コマ、カラーシートフィルム4,350枚のデジタル化を完了した。

史料収集時の諸情報、一点目録情報で構成されるメタデータとともに、デジタル化した史料画像やデジタル撮影データを格納するアーカイヴハブを構築した。

キーワード検索・撮影年の範囲指定検索

による画像検索システムを構築し、史料編纂所閲覧室においてマイクロフィルム簡易検索として公開した。

画像閲覧システムを構築し、紙媒体の「写真帳」では望みえなかった画像の拡大縮小表示や、紙背文字を読み易くする上下・左右反転表示など一般的ビューワでは実現しにくい機能を実装した。

厳密な制御が不可欠な閲覧・印刷に関し、Web コンテンツ保護システムを研究開発し、閲覧・コピー&ペースト・印刷・画面キャプチャーのグループ毎の制限を実現した。

採訪から利用までの一貫した管理を可能としたデジタル撮影による史料収集の仕様を策定し、これに基づくデジタル撮影画像をアーカイヴハブに格納し公開に供した。

メタデータ検索と画像閲覧を実現するため、一史料画像毎に差出(発信者)・宛所(受信者)・発信年代・内容等を研究して史料名を付与したデジタル史料画像約100万件余りの一点目録を作成した。

一点目録検索システムとして、史料編纂所所蔵史料目録データベース(Hi CAT)と連携したHi CAT Plusを開発し、史料編纂所閲覧室において公開した。

Hi CAT との横断検索を実現し、厳密なセキュリティコントロールを持つHi CAT Plus 経由による各データベースとデジタル史料画像群との連携を可能とした。

アーカイヴハブ/Hi-CAT Plus システムに、歴史=史料知識オントロジー機能を搭載するため、史料群の所蔵先(現在)の地域毎の検索(キーワード・都道府県・旧国郡)が可能なる「史料所在オントロジーシステム」を開発・実装した。また、デジタル画像のサムネイル生成に伴って極度な遅延化が生じたアップロードツールを改善した。

秋田藩家蔵文書デジタル収集とバーチャル「秋田藩採集文書」の構築に関する秋田県公文書館との連携システムを実現した。

SHIPS 古文書フルテキストデータベース上に奥野高廣氏編『増訂織田信長文書の研究』(吉川弘文館)上巻・下巻・補遺に収録される文書全文データを追加した。

萩原寺地蔵院聖教プロジェクト・金光図書館具注暦プロジェクトに関して、それぞれのデジタル化・目録作成を進め、研究成果報告書として刊行した。

電子くずし字字典データベースを軸とした機関間連携の推進、および字形・字体データから人物オントロジーへの展開に関する研究に関して、奈良文化財研究所「木簡画像データベース・木簡字典」との連携システムをより一層推進した。

他の外部資金も獲得し古写真プロジェクト・中国档案プロジェクトを進めた。

(3)2013年2月28日、東京大学史料編纂所において、研究期間5年間の研究を総括する公開研究集会「S 科研「史料デジタル収集の体系化に基づく歴史オントロジー構築の研究」の成果と課題」を開催・報告した。

(4)これらの成果を踏まえて、2013年6月28日に、研究成果報告書『史料デジタル収集の体系化に基づく歴史オントロジー構築の研究』(東京大学史料編纂所研究成果報告書2013-2)、ならびに同『金光図書館の具注暦』(東京大学史料編纂所研究成果報告書2013-1)を刊行し、同年9月18日に、本研究成果の一部である同『佐竹義宣書状集 梅津憲忠宛』(東京大学史料編纂所研究成果報告書2013-3)を刊行した。

5. 主な発表論文等

〔雑誌論文〕(計48件)

井上聡「『電子くずし字字典データベース』の課題と将来構想」、『人文科学とコンピュータ研究会報告』2013-CH-97、No.10、有、2013、pp.1~7

金子拓「肥後加藤家旧蔵豊臣秀吉・秀次朱印状について(続)」、『東京大学史料編纂

所研究紀要』、無、22号、2012、pp.82~101
林讓「東京大学史料編纂所における研究資源の利活用と保全について」、『人間文化研究情報資源共有化研究会報告集』、無、3、2012、pp.49~53
大内英範・山田太造・高橋典幸・綱川歩美・林讓・保谷徹・山家浩樹・横山伊徳「Hi-CAT Plus: デジタル史料の検索・閲覧システム」、『人文科学とコンピュータシンポジウム論文集』、有、2011-8、2011、pp.105~110
山田太造・横山伊徳・綱川歩美・高橋典幸・林讓「採訪史料管理システム」、『じんもんこん 2010』、有、2010、pp.145~150
横山伊徳「史料編纂とデジタル化のメタヒストリー」、『人工知能学会誌』、有、25-1号、2010、pp.5~10
横山伊徳「日本史研究データベースはどこへ行こうとしているのか」、『日本歴史』、有、740号、2010、pp.2~8
金子拓「秋田藩家蔵文書のデータベース化と地域連携 「秋田藩採集文書論」のために」、『東京大学史料編纂所シンポジウム研究と情報の資源化』予稿集、無、2009、pp.5~25
石川徹也・伊藤直之・前澤克俊「編纂史料の高度検索システム構築研究と史料のデジタル化」、『画像電子学会第36回年次大会』予稿集、無、2008、pp.1~9
石川徹也「歴史知識 Ontology 構築研究 その論理と実際」、『情報文化学会誌』、有、15巻1号、2008、pp.6~11
石川徹也・北内啓・城塚音也「歴史オントロジー構築のための史料からの人物情報抽出」、『自然言語処理』、有、15巻4号、2008、pp.3~18

[学会発表](計75件)

林讓「S科研の課題と経過」/綱川歩美「デジタルデータ化の基礎と方法」/大内英範「Hi-CAT Plus の公開: セキュリティ、原本所蔵者との共有システム」/山田太造

「デジタル化した収集史料の管理・検索システム」/横山伊徳「オントロジーの発想と史料所在知識の表現」/金子拓「S科研による史料のデジタル収集と調査研究・データベース化の成果」/太田研・三浦貴之「秋田県公文書館におけるデジタルアーカイブの取組み」/山家浩樹「史料デジタル収集の基準と方法」/保谷徹「S科研の今後の課題」、公開研究会『「S科研」史料デジタル収集の体系化に基づく歴史オントロジー構築の研究』の成果と課題』、2013年2月28日、於東京大学史料編纂所
大内英範・山田太造・高橋典幸・綱川歩美・林讓・保谷徹・山家浩樹・横山伊徳「Hi-Cat Plus: デジタル史料の検索・閲覧システム」、『じんもんこん 2011』、2011年12月10日、於龍谷大学
林讓「デジタル史料画像の検索・閲覧システムについて」、日本古文書学会学術大会、2011年9月25日、於国学院大学
大内英範「史料編纂所の公開 DB とデジタルアーカイブ(Databases and the Digital Archive of the Historiographical Institute)」、『European Association of Japanese Resource Specialists (日本資料専門家欧州協会)』、2011年9月8日、於Newcastle 大学(イギリス)
山田太造・横山伊徳・大内英範・綱川歩美・高橋典幸・林讓「デジタル史料写真帳: 収集史料のデジタル化と検索・閲覧システム」、2011年度画像電子学会第39回年次大会、2011年6月25日、於くにびきメッセ
山田太造・横山伊徳・綱川歩美・高橋典幸・林讓「採訪史料管理システム」、『じんもんこん 2010』(<http://www.jinmoncom.jp/sympo2010/>)、2010年12月12日、於東京工業大学
金子拓「秋田藩家蔵文書のデータベース化と地域連携 - 「秋田藩採集文書論」 - 」/林讓「科学研究費基盤研究(S)「史料デジ

タル収集の体系化に基づく歴史オントロジー構築の研究」の現状と課題について」公開シンポジウム「史料研究と情報の資源化 - 史料編纂所大型プロジェクトの進捗」、2010年1月30日、於東京大学山上会館
石川徹也・横山伊徳・赤石美奈・遠藤基郎・近藤成一「歴史知識学の創成」、東京大学史料編纂所附属前近代日本史情報国際センター公開研究会、2008年11月22日、於東京大学山上会館

〔図書〕(計55件)

及川亘・加藤昌宏・金子拓『佐竹義宣書状集 梅津憲忠宛』、2013、168

林讓・網川歩美・山家浩樹・山田太造・太内英範・高橋典幸・横山伊徳・金子拓・太田研・三浦貴之・厚谷和雄・井上聡・保谷徹『史料デジタル収集の体系化に基づく歴史オントロジー構築の研究』、2013、145

厚谷和雄・末柄豊・伴瀬明美・藤原重雄『金光図書館の具注暦』、2013、102

西脇康『佐渡小判・切銀の研究 付、佐渡銀』、書信館出版株式会社、2013、252

石川徹也・横山伊徳『つながる図書館・博物館・文書館 デジタル化時代の知の基盤づくりへ』、東京大学出版会、2011、272

村井祐樹・末柄豊『真如寺所蔵能勢家文書』、2010、70

石川徹也・横山伊徳・赤石美奈・遠藤基郎・近藤成一『歴史知識学ことはじめ』、勉誠出版、2009、202

〔その他〕

ホームページ等

東京大学史料編纂所共同研究「2012年度史料デジタル収集の体系化に基づく歴史オントロジー構築の研究」<http://www.hi.u-tokyo.ac.jp/collaboration/kaken/2022001.html>、2008～2011年度の概要も同様。

新聞掲載・TV放映

(1)電子くずし字字典データベース機関間

連携プロジェクトによる国立文化財機構奈良文化財研究所「木簡画像データベース・木簡字典」との連携検索：2009年5月30日付奈良日日新聞・奈良新聞・産経新聞・毎日新聞・朝日新聞・読売新聞朝刊、2009年6月10日付文教速報、および2009年10月15日付NIKKEI NET・奈良新聞WEB・YOMIURI ON LINE・産経関西WEB、同14日付奈良テレビ放送、同15日付NHK奈良
(2)金光図書館具注暦プロジェクトによる金光図書館所蔵史料調査：2010年3月27日付山陽新聞

公開研究会等(上記記載以外)

WEBコンテンツ保護勉強会(2010年1月26日、曾根原登氏「Web情報流通の課題と動向」、越前功氏「デジタルコンテンツの保護技術」、柿沼澄男氏「共同利用・共同研究拠点と大学共同利用機関」)

6. 研究組織

(1)研究代表者

林讓(HAYASHI, Yuzuru)
東京大学・史料編纂所・教授
研究者番号：00164971

(2)研究分担者

横山伊徳(YOKOYAMA, Yoshinori)
東京大学・史料編纂所・教授
研究者番号：90143536

加藤友康(KATO, Tomoyasu)
明治大学・大学院・特任教授
研究者番号：00114439
(平成20～21年度)

保谷徹(HOYA, Toru)
東京大学・史料編纂所・教授
研究者番号：60195518

久留島典子(KURUSHIMA, Noriko)

東京大学・史料編纂所・教授
研究者番号：70143534

山家 浩樹 (YANBE, Koki)
東京大学・史料編纂所・教授
研究者番号：60191467

(3)連携研究者

石川 徹也 (ISHIKAWA, Tetsuya)
東京大学・史料編纂所・特任教授
研究者番号：20041808

井上 聡 (INOUE, Satoshi)
東京大学・史料編纂所・助教
研究者番号：20302656

榎原 雅治 (EBARA, Masaharu)
東京大学・史料編纂所・教授
研究者番号：40160379

遠藤 基郎 (ENDO, Motoo)
東京大学・史料編纂所・准教授
研究者番号：40251475

大内 英範 (OOUCHI, Hidenori)
東京大学・史料編纂所・特任助教
研究者番号：60462173

尾上 陽介 (ONOE, Yosuke)
東京大学・史料編纂所・准教授
研究者番号：00242157

金子 拓 (KANEKO, Hiraku)
東京大学・史料編纂所・助教
研究者番号：10302655

木村 直樹 (KIMURA, Naoki)
東京大学・史料編纂所・助教
研究者番号：40323662

小宮 木代良 (KOMIYA, Kiyora)
東京大学・史料編纂所・教授
研究者番号：90186809

近藤 成一 (KONDO, Shigekazu)
東京大学・史料編纂所・教授
研究者番号：90153717

末柄 豊 (SUEGARA, Yutaka)
東京大学・史料編纂所・准教授
研究者番号：70251478

藤原 重雄 (FUJIWARA, Shigeo)
東京大学・史料編纂所・助教
研究者番号：40313192

松澤 克行 (MATSUZAWA, Yoshiyuki)
東京大学・史料編纂所・准教授
研究者番号：40282529

山田 太造 (YAMADA, Taizo)
人間文化研究機構本部・特任研究員
研究者番号：70413937

赤石 美奈 (AKAISHI, Mina)
法政大学・情報科学部・教授
研究者番号：60273166

黒田 日出男 (KURODA, Hideo)
立正大学・文学部・教授
研究者番号：90013284

高橋 典幸 (TAKAHASHI, Noriyuki)
東京大学大学院・人文社会研究科・准教授
研究者番号：10292799

石川 寛夫 (ISHIKAWA, Hiroo)
九州産業大学・芸術学部・前教授
研究者番号：40104864
(平成20~23年度)